The Theorems of Sard and Whitney

Swayam Chube

Last Updated: May 10, 2025

Contents

§1 Sard's Theorem

We begin by proving Sard's theorem for smooth functions between Euclidean spaces. As is customary in differential topology, it will be quite easy to port these results to smooth maps between arbitrary differentiable manifolds. The following is taken from [Mil97].

THEOREM 1.1. Let $F\colon U(\subseteq \mathbb{R}^n) \to \mathbb{R}^p$ be a smooth map, with U open in $\mathbb{R}^n$. Let $C \subseteq U$ be the set of critical points of F. Then $F(C) \subseteq \mathbb{R}^p$ has measure zero.

Proof. The proof will be by induction on $n \geqslant 0$. The base case with $n = 0$ and $p \geqslant 1$ is trivial. Define

$$C_i := \{x \in U : \text{all partial derivatives of order} \leqslant i \text{ vanish at } x\} \subseteq C.$$

This gives us a descending chain of closed subsets of U:

$$C \supseteq C_1 \supseteq C_2 \supseteq C_3 \supseteq \cdots.$$

We shall prove the theorem in three steps.
STEP 1. *The image $F(C \setminus C_1)$ has measure zero.*

We may assume that $p \geqslant 2$, since $C = C_1$ when $p = 1$. For each $x \in C \setminus C_1$, we will find an open neighborhood $V \subseteq U$ so tht $F(V \cap C)$ has measure zero. Since $C \setminus C_1$ is covered by countably many of these neighborhoods, this will prove that $F(C \setminus C_1)$ has measure zero.

Since $x \notin C_1$, there is some partial derivative which is not zero at x. We may suppose without loss of generality that $\partial F_1/\partial x_1$ is non-zero at x. Consider the map $h\colon U \to \mathbb{R}^n$ defined by

$$h(y) = (F_1(y), y_2, \ldots, y_n).$$

The Jacobian matrix of the above function at x is clearly non-singular, and hence, h maps some neighborhood V of x contained in U diffeomorphically onto an open set V'. The composition $G =$

$F \circ h^{-1}: V' \to \mathbb{R}^p$ will then map V' into $\mathbb{R}^p$. Note that the set C' of critical points of G is precisely $h(V \cap C)$, and hence, the set $G(C')$ of critical values of G is equal to $F(V \cap C)$.

For each $(t, y_2, \ldots, y_n) \in V'$, note that $G(t, y_2, \ldots, y_n)$ belongs to the hyperplane $\{t\} \times \mathbb{R}^{p-1} \subseteq \mathbb{R}^p$: thus G carries hyperplanes into hyperplanes. Let

$$G^t: \left(\{t\} \times \mathbb{R}^{p-1}\right) \cap V' \to \{t\} \times \mathbb{R}^{p-1}$$

denote the restriction of G. Note that a point of $\{t\} \times \mathbb{R}^{n-1}$ is critical for G^t if and only if it is critical for G, since the Jacobian matrix of G has the form

$$\left(\frac{\partial G_i}{\partial x_j}\right) = \begin{pmatrix} 1 & 0 \\ * & \left(\frac{\partial G_i^t}{\partial x_j}\right) \end{pmatrix}.$$

According to the induction hypothesis, the set of critical values of G^t has measure zero in $\{t\} \times \mathbb{R}^{p-1}$. Therefore the set of critical values of G intersects each hyperplane $\{t\} \times \mathbb{R}^{p-1}$ in a set of measure zero. Hence, by Fubini's theorem, the set $G(C') = F(V \cap C)$ has measure zero in $\mathbb{R}^p$, thereby completing the proof of this step.

Step 2. *The image $F(C_i \setminus C_{i+1})$ has measure zero for $i \geqslant 1$.*

For each $x \in C_k \setminus C_{k+1}$, there is some $(k+1)$-st derivative

$$\frac{\partial^{k+1} F_r}{\partial x_{s_1} \cdots \partial x_{s_{k+1}}}$$

which is non-zero at x. Thus the function

$$w = \frac{\partial^k F_r}{\partial x_{s_2} \cdots \partial s_{k+1}}$$

vanishes at x but $\partial w / \partial x_{s_1}$ does not. Suppose, without loss of generality that $s_1 = 1$. Then the map $h: U \to \mathbb{R}^n$ defined by

$$h(y) = (w(y), y_2, \ldots, y_n)$$

has non-singular Jacobian matrix at $y = x$ and thus carries some neighborhood V of x contained in U diffeomorphically onto an open set V' of $\mathbb{R}^n$. Note that h carries $C_k \cap V$ into the hyperplane $\{0\} \times \mathbb{R}^{n-1}$. Set

$$G = F \circ h^{-1}: V' \to \mathbb{R}^p$$

and let

$$\overline{G}: \left(\{0\} \times \mathbb{R}^{n-1}\right) \cap V' \to \mathbb{R}^p$$

denote the restriction of G. By induction, the set of critical values o $\overline{G}$ has measure zero in $\mathbb{R}^p$. But each point in $h(C_k \cap V)$ is clearly a critical point of $\overline{G}$, since all derivatives of order $\leqslant k$ vanish. Therefore

$$\overline{G} \circ h(C_k \cap V) = F(C_k \cap V)$$

has measure zero in $\mathbb{R}^p$. Since $C_k \setminus C_{k+1}$ is covered by countably many such sets V, it follows that $F(C_k \setminus C_{k+1})$ has measure zero in $\mathbb{R}^p$.

Step 3. *The image $F(C_k)$ has measure zero for $k \gg 0$.*

Let $I^n \subseteq U$ be a cube with edge δ. If k is sufficiently large, we shall show that $F(C_k \cap I^n)$ has measure zero. Since C_k can be covered by countably many such cubes, this will prove that $F(C_k)$ has measure zero.

Using Taylor's theorem, the compactness of I^n, and the definition of C_k, we see that

$$F(x+h) = F(x) + R(x,h), \tag{1}$$

where

$$\|R(x,h)\| \leqslant c\|h\|^{k+1}$$

for $x \in C_k \cap I^n$ and $x+h \in I^n$, where c is a constant which depends only on F and the cube I^n. For a positive integer $r > 0$, subdivide I^n into r^n cubes of edge length δ/r. Let I_1 be a cube of the subdivision which contains a point x of C_k. Then any point of I_1 can be written as $x+h$ with $\|h\| \leqslant \sqrt{n}(\delta/r)$. Hence, from (1), it follows that $F(I_1)$ is contained in a cube of edge length a/r^{k+1} centered about $F(x)$, where $a = 2c\left(\sqrt{n}\delta\right)^{k+1}$ is a constant. Hence $F(C_k \cap I^n)$ is contained in a union of at most r^n cubes having total volume

$$\text{Volume} \leqslant r^n \left(\frac{a}{r^{k+1}}\right)^p = a^p r^{n-(k+1)p}.$$

For sufficiently large k, it is clear that the above volume tends to zero as $r \to \infty$, so $F(C_k \cap I^n)$ must have measure zero, thereby completing the proof of the theorem. ■

Next, we define the notion of "measure zero" on an arbitrary manifold.

Definition 1.2. A subset A of a smooth n-manifold is said to have *measure zero* if for every smooth chart (U,φ) of M, the set $\varphi(A \cap U)$ has measure zero in $\mathbb{R}^n$.

Lemma 1.3. Let $A \subseteq \mathbb{R}^n$ have measure zero and $F\colon A \to \mathbb{R}^n$ be a smooth map. Then $F(A) \subseteq \mathbb{R}^n$ has measure zero.

Proof. By definition, for each point $p \in A$, there is a neighborhood U of p on which F extends to a smooth function. Choose a compact ball $\overline{B}$ contained in U. Note that A can be covered by countably many such balls, and as such, it suffices to show that $F(A \cap \overline{B})$ has measure zero in $\mathbb{R}^n$.

Since $\overline{B}$ is compact, the (Frobenius) norm of the Jacobian matrix

$$DF = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{pmatrix}$$

is bounded by some constant $C > 0$. Then, for $x, y \in \overline{B}$, we have, using the mean value theorem that

$$\begin{aligned} \|F(y) - F(x)\| &= \left\| \int_0^1 \frac{d}{dt} F(ty + (1-t)x)\, dt \right\| \\ &\leqslant \int_0^1 \|DF(ty + (1-t)x)\| \|y - x\|\, dt \\ &\leqslant C\|y - x\|. \end{aligned}$$

Now, given $\delta > 0$, we can choose a countable cover $\{B_j\}$ of $A \cap \overline{B}$ by open balls satisfying

$$\sum_j \mu(B_j) < \delta,$$

where μ denotes the standard Lebesgue measure on $\mathbb{R}^n$. Then due to our computations above, $F(\overline{B} \cap B_j)$ is contained in a ball $\widetilde{B}_j$ whose diameter is no more than C times that of B_j. In particular, this means that

$$\sum_j \mu(\widetilde{B}_j) < C^n \delta.$$

Since this quantity can be made as small as desired, it follows that $F(A \cap \overline{B})$ has measure zero in $\mathbb{R}^n$, thereby completing the proof. ■

LEMMA 1.4. Suppose A is a subset of a smooth n-manifold M, and for some collection $\{(U_\alpha, \varphi_\alpha)\}$ of smooth charts whose domains cover A, $\varphi_\alpha(A \cap U_\alpha)$ has measure zero in $\mathbb{R}^n$ for each α. Then A has measure zero in M.

Proof. Let (V, ψ) be an arbitrary smooth chart. We would like to show that $\psi(A \cap V)$ has measure zero in $\mathbb{R}^n$. Since the manifold is second countable, there is a countable subset $\{U_\beta\}$ of $\{U_\alpha\}$ which covers V. Note that

$$\psi(A \cap V) = \bigcup_\beta (\psi \circ \varphi_\beta^{-1}) \left[\varphi_\beta \left(A \cap V \cap U_\beta\right)\right],$$

which is clearly measure zero due to Lemma 1.3 ■

THEOREM 1.5 (SARD). Let $F \colon M \to N$ be a smooth map between manifolds. Then the set of critical values of F in N has measure zero.

Proof. Follows immediately from Theorem 1.1 and the discussion above. ■

COROLLARY 1.6. Let $m < n$, M a smooth m-manifold and N a smooth n-manifold. If $F \colon M \to N$ is a smooth map, then the image of F is of measure zero in N.

Proof. The image of F is precisely the set of critical values in N. ■

§2 Whitney's Theorems

§§ The Immersion and Embedding Theorems

THEOREM 2.1. Let $F \colon M \to \mathbb{R}^m$ be any smooth map, where M is a smooth n-manifold, and $m \geqslant 2n$. For any $\varepsilon > 0$, there is a smooth immersion $\widetilde{F} \colon M \to \mathbb{R}^m$ such that

$$\sup_M |\widetilde{F} - F| \leqslant \varepsilon.$$

Proof. Let $\{W_i\}$ be a regular[1] open cover of M, which exists due to [Lee03, Proposition 2.24]. Then each W_i is the domain of a smooth chart $\psi_i \colon W_i \to B_3(0)$ and the precompact sets $U_i = \psi_i^{-1}(B_1(0))$ still cover M. For each integer $k \geqslant 1$, set $M_k = \bigcup_{i=1}^k U_i$ which is an open submanifold of M. We interpret M_0 to be the empty set. We shall modify F inductively on one set W_i at a time.

For each $i \geqslant 1$, let $\varphi_i \in C^\infty(M)$ be a smooth bump function supported in W_i tht is equal to 1 on $\overline{U}_i$. Let $F_0 = F$, and suppose by induction we have defined smooth maps $F_j \colon M \to \mathbb{R}^m$ for $0 \leqslant j \leqslant k-1$ satisfying

(i) $\sup_M |F_j - F| \leqslant \varepsilon$;

(ii) If $j \geqslant 1$, $F_j(x) = F_{j-1}(x)$ for all $x \in M \setminus W_j$;

[1] An open cover $\{W_i\}$ of M is said to be *regular* if it satisfies the following properties:

- The cover $\{W_i\}$ is countable and locally finite;
- Each W_i is the domain of a smooth coordinate map $\varphi_i \colon W_i \to B_3(0) \subseteq \mathbb{R}^n$;
- The collection $\{U_i\}$ still covers M, where $U_i = \varphi^{-1}(B_1(0))$.

(iii) The differential $(F_j)_*$ is injective at each point of $\overline{M}_j$.

For any $m \times n$ matrix A, define a new map $F_A : M \to \mathbb{R}^m$ as follows: On $M \setminus \operatorname{Supp}\varphi_k$, $F_A = F_{k-1}$, and on W_k, F_A is the map given in coordinates (through the chart ψ_k) by

$$F_A(x) = F_{k-1}(x) + \varphi_k(x)Ax,$$

i.e., for $\xi \in W_k$, set $x = \psi_k(\xi)$ in the above equation. Since both definitions agree on the set $W_k \setminus \operatorname{Supp}\varphi_k$, the map F_A is smooth.

Because (i) holds for $j = k-1$, there is a constant $\varepsilon_0 < \varepsilon$ such that $|F_{k-1}(x) - F(x)| \leqslant \varepsilon_0$ on the compact set $\operatorname{Supp}\varphi_k$. By continuity, there is some $\delta > 0$ such that

$$\sup_M |F_A - F_{k-1}| = \sup_{x \in \operatorname{Supp}\varphi_k} \left|\varphi_k(x)Ax\right| < \varepsilon - \varepsilon_0,$$

and therefore,

$$\sup_M |F_A - F| \leqslant \sup_M |F_A - F_{k-1}| + \sup_M |F_{k-1} - F| < \varepsilon.$$

Let $P : W_k \times \mathbb{M}(m \times n, \mathbb{R}) \to \mathbb{M}(m \times n, \mathbb{R})$ be the matrix-valued function

$$P(x, A) = DF_A(x).$$

According to the inductive hypothesis, the matrix $P(x, A)$ has rank n when (x, A) is in the compact set $\left(\operatorname{Supp}\varphi_k \cap \overline{M}_{k-1}\right) \times \{0\}$. Again, by choosing δ even smaller if necessary, we may also ensure that $\operatorname{rank} P(x, A) = n$ whenever $x \in \operatorname{Supp}\varphi_k \cap M_{k-1}$ and $|A| < \delta$[2].

Finally, we need to ensure that $\operatorname{rank}(F_A)_* = n$ on $\overline{U}_k$ and therefore on $\overline{M}_k = \overline{M}_{k-1} \cup \overline{U}_k$. Note that

$$DF_A(x) = DF_{k-1}(x) + A$$

for $x \in \overline{U}_k$ because $\varphi_k \equiv 1$ on that set. Hence, $DF_A(x)$ has rank n in $\overline{U}_k$ if and only if A is not of the form $B - DF_{k-1}(x)$ or any $x \in \overline{U}_k$ and any matrix $B \in \mathbb{M}(m \times n, \mathbb{R})$ of rank less than n.

To this end, let $Q : W_k \times \mathbb{M}(m \times n, \mathbb{R}) \to \mathbb{M}(m \times n, \mathbb{R})$ be the smooth map

$$Q(x, B) = B - DF_{k-1}(x).$$

We need to show that there is some matrix A with $|A| < \delta$ that is not of the form $Q(x, B)$ for any $x \in \overline{U}_k$ and any matrix $B \in \mathbb{M}(m \times n, \mathbb{R})$ of rank less than n.

For $0 \leqslant j \leqslant n-1$, the set $\mathbb{M}_j(m \times n, \mathbb{R})$ of $m \times n$ matrices of rank j is an embedded submanifold of $\mathbb{M}(m \times n, \mathbb{R})$ of codimension $(m-j)(n-j)$[3]. Due to Corollary 1.6, the image under Q of $W_k \times \mathbb{M}_j(m \times n, \mathbb{R})$ has measure zero provided that the dimension of $W_k \times \mathbb{M}_j(m \times n, \mathbb{R})$ is strictly less than that of $\mathbb{M}(m \times n, \mathbb{R})$, i.e.,

$$n + mn - (m-j)(n-j) < mn \iff n < (m-j)(n-j).$$

For $j < n$, note that $(m-j) > n$ and $n - j \geqslant 1$, therefore, $n < (m-j)(n-j)$. Thus, the image of each $W_k \times \mathbb{M}_j(m \times n, \mathbb{R})$ under Q has measure zero in $\mathbb{M}(m \times n, \mathbb{R})$. Thus, choosing $A \in \mathbb{M}(m \times n, \mathbb{R})$ with $|A| < \delta$ and not in the union of those image sets, and setting $F_k = F_A$, we obtain a map satisfying (i), (ii), and (iii) for $j = k$.

Now let

$$\widetilde{F}(x) = \lim_{k \to \infty} F_k(x).$$

[2] This follows from a "tube lemma" type argument.

[3] See [Lee03, Example 8.14].

Note that due to the local finiteness of the cover $\{W_i\}$, for each k, the sequence $\{\widetilde{F}_k(x)\}$ is locally constant and equal to some F_N in a neighborhood of each point. In particular, the limit $\widetilde{F}$ clearly exists and is smooth because it agrees locally with smooth functions. Furthermore, due to (iii), $\widetilde{F}$ is an immersion. ■

COROLLARY 2.2 (WHITNEY IMMERSION THEOREM). Every smooth n-manifold admits an immersion into $\mathbb{R}^{2n}$.

Proof. Apply Theorem 2.1 to the constant map $F \equiv 0 \colon M \to \mathbb{R}^{2n}$. ■

THEOREM 2.3. Let $F \colon M \to \mathbb{R}^m$ be a smooth immersion, where M is a smooth n-manifold, and $m \geqslant 2n+1$. Then for any $\varepsilon > 0$ there is an injective immersion $\widetilde{F} \colon M \to \mathbb{R}^m$ such that

$$\sup_M |\widetilde{F} - F| \leqslant \varepsilon.$$

§§ The Approximation Theorems

§§ Tubular Neighborhoods

§§ Smooth Approximations of Continuous Functions

References

[Lee03] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2003.

[Mil97] John W. Milnor. *Topology from the differentiable viewpoint*. Princeton University Press, 1997.